Journal of Hydrology 570 (2019) 343–355

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts

Yanlai Zhou[a,b,c], Shenglian Guo[a], Fi-John Chang[b,*]

[a] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
[b] Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan
[c] Department of Geosciences, University of Oslo, Oslo 0316, Norway

ARTICLE INFO

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Shreedhar Maskey, Associate Editor

Keywords:
Artificial Intelligence (AI)
Recurrent ANFIS
Evolutionary algorithm
Multi-step-ahead flood forecast
Time series
Three Gorges Reservoir (TGR)

ABSTRACT

Reliable and precise multi-step-ahead flood forecasts are crucial and beneficial to decision makers for mitigating flooding risks. For a river basin undergoing fast urban development, its regional meteorological condition interacts frequently with intensive human activities and climate change, which gives rise to the non-stationary process between rainfall and runoff whose non-stationary features is difficult to be captured by a non-recurrent data-driven model with a static learning mechanism. This study proposes a recurrent Adaptive-Network-based Fuzzy Inference System (R-ANFIS) embedded with Genetic Algorithm and Least Square Estimator (GL) that optimize model parameters for making multi-step-ahead forecasts. The main merit of the proposed method (R-ANFIS(GL)) lies in capturing the features of the non-stationary process between rainfall and runoff series as well as in alleviating time-lag effects encountered in multi-step-ahead flood forecasting. To demonstrate model reliability and effectiveness, the R-ANFIS(GL) model was implemented to make multi-step-ahead forecasts from horizons t + 1 up to t + 8 for a famous benchmark chaotic time series and a flood inflow series of the Three Gorges Reservoir (TGR) in China. For comparison purpose, two ANFIS neural networks of different structures (one dynamic and one static neural networks) were also implemented. Numerical and experimental results indicated that the R-ANFIS(GL) model not only outperformed the two comparative networks but significantly enhanced the accuracy of multi-step-ahead forecasts for both chaotic time series and the reservoir inflow case during flood seasons, where effective mitigation of time-lag bottlenecks was achieved. We demonstrated that the R-ANFIS(GL) model could suitably configure the complex non-stationary rainfall-runoff process and effectively integrate the monitored rainfall and discharge data with the latest outputs of the model so that the time shift problem could be alleviated and model reliability as well as forecast accuracy for future horizons could be significantly improved.

1. Introduction

A reliable and accurate long-term flood forecast model yields minimal error, and it facilitates decision-making on the optimal reservoir operation for achieving minimal flood risks and/or maximal operational synergies as well as allowing sufficient time to prepare for hazard management. This is particularly important for real-time reservoir operation during flood periods. Nevertheless, rainfall and runoff variables are both spatially and temporally fickle while notoriously interrelated such that rendering the watershed system is a highly complex, dynamic and non-stationary process, which is very challenging. Moreover, it is much difficult to predict the multi-step, rather than single-step, chaotic nature in real-world time series due to the high uncertainty in inputs and the interactions between different prediction horizons. Multi-step-ahead forecasting has been carried out mostly with recursive strategies (Taieb and Atiya, 2016). The model input selection (e.g., data pre-processing using data-mining techniques), model parameter optimization and model output post-processing (e.g., ensemble forecasts, real time correction) are the main foci and important components in multi-step-ahead hydrological forecasts. In this study, we pay special attention to model parameter optimization. It is a valuable strategy to use recurrent mechanisms that adopt model outputs as external (extra) information for making real-time multi-step-ahead forecasts, whereas local search algorithms may bring instability and local minima bottlenecks that would easily give rise to error accumulation and propagation (Abrahart et al. 2012; Banihabib et al. 2015; Chang et al., 2014; Nanda et al. 2016; Tran et al. 2016; Zhang et al. 2018a). In addition, such time-lag problems may cause significant performance

* Corresponding author.
E-mail address: changfj@ntu.edu.tw (F.-J. Chang).

https://doi.org/10.1016/j.jhydrol.2018.12.040
Received 21 May 2018; Received in revised form 15 November 2018; Accepted 10 December 2018
Available online 14 January 2019

deterioration when coping with multi-step-ahead forecasts in real-world applications. When making multi-step-ahead streamflow forecasts during flood seasons, models with time-lag problems (e.g. model overfitting/instability or real-time observed rainfall/runoff values unavailable) would fail to trace flow trails closely, especially peak flows, as forecast horizon increases. Regarding multi-step-ahead forecasting, it has been argued for a long time on whether global search algorithms that optimize model parameters for recurrent mechanisms, like anterior observations and/or predicted values, could overcome the technical drawbacks like time-lag phenomena of local search algorithms. Consequently, an efficient algorithm is required for determining the optimal network parameter setting to improve the stability and reliability of forecast models.

Artificial neural networks (ANNs) and fuzzy inference systems (FIS) are crucial branches of Artificial Intelligence (AI) techniques, and they have been successfully utilized to model various hydrological time series in the last two decades (e.g., Chen et al. 2013; He et al., 2014; Nourani et al. 2014; Fazel et al., 2015; Chang and Tsai 2016; Alexander and Thampi, 2018; Fotovatikhah et al., 2018; Tan et al., 2018). The main characteristic of an ANN is to adapt itself automatically to the simulated environment through a large amount of input–output patterns (Kumar et al. 2015; Kasiviswanathan et al. 2016; Shoaib et al. 2016; Taormina et al., 2015; Wang et al., 2013; Zhang et al. 2018b). This adaptation allows taking account of the specific features of each investigative event to keep model forecast capability at the average in the long run. The imprecision and uncertainty of model inputs could be handled by the FIS, while the input–output patterns could be identified by ANNs with adaptive learning capabilities (Yaseen et al., 2015, 2017). The Adaptive-Network-based Fuzzy Inference System (ANFIS) proposed by Jang (1993) integrates adaptive learning capability with fuzzy reasoning and is capable of coping with complexity as well as noise problems, such as streamflow forecasts (Chang and Chang, 2006; Firat and Güngör, 2008; He et al., 2014; Tsai et al., 2014). However, rainfall-runoff is a time-dependent dynamic process with input-output patterns difficult to identify entirely using the static ANFIS. Lately, a considerable amount of studies has migrated to explore recurrent fuzzy neural networks for analyzing time series and temporal processes (Kasabov and Song 2002; Nguyen et al., 2018). The recurrent ANFIS (R-ANFIS) with a dynamic mechanism has feedback connections in its topology, where current external variables and delayed outputs constitute the inputs of the recurrent model (Zhang and Morris 1999). Such recurrent mechanism possesses prominent capability that the mapping between inputs and outputs remains dynamic, rather than stationary, over time and long-short term memories can be incorporated into the ANFIS when making multi-step-ahead time series forecasts (e.g., Mastorocostas and Theocharis 2002; Fei and Lu 2018; Xiong and Zhang 2018). Such characteristic and ability of the R-ANFIS, however, does not receive much attention from environmental and hydrological fields.

Recurrent networks are usually trained by the Gradient Descent Algorithm (GDA). The local search algorithm proposed by Jang (1993) for the ANFIS is a combination of the Least Square Estimator (LSE) and the Steepest Descent Algorithm (SDA) such that linear and nonlinear parameters are locally adjusted to minimize the errors between outputs and targets. Both the SDA and the GDA perform derivative operation, which usually encounters instability as well as local minima problems during modelling (Tamura et al. 2008). Evolutionary algorithms are promising tools to accomplish the parametric learning of ANNs. The main feature of these algorithms reveals that the locations of the extrema of a function defined over the search space depend on a population of positions, instead of a single position, in the search space (Chandra, 2015; Liu et al. 2016). In comparison with other evolutionary algorithms, such as the Particle Swarm Optimization (PSO) and the Differential Evolution (DE) algorithms, the Genetic Algorithm (GA) proposed by Holland, 1975 has the ability to mimic processes observed in natural evolution and is prevalently utilized to solve various optimization problems in hydrology and water resources fields (e.g., Cheng et al., 2005; Chang et al., 2010; Chen and Chang, 2009; Zhou and Guo, 2013; Li et al. 2014; Naghibi, et al., 2017; Yin et al., 2017; Zhou, et al. 2017; Ehteram, et al., 2018).

The innovative nature of this study lies in the application of the R-ANFIS(GL) for the first time in multi-step-ahead flood forecasting, and we propose to adapt fuzzy learning equipped with evolutionary parameter optimization by means of predictive recurrence from antecedent prediction for future prediction horizons. We expect to effectively mitigate time-lag effects caused by the local search algorithm and significantly increase prediction accuracy. To demonstrate the reliability and applicability of the proposed approach in multi-step-ahead flood forecasts, we utilize the Mackey-Glass time series as a benchmark and the inflow series of the Three Gorges Reservoir in China as a real application case. Needless to say, the famous Mackey-Glass and the Lorenz time series are two of the most referred benchmark nonlinear chaotic time series. The former is adopted as the benchmark in this study, inasmuch as the former generated by a nonlinear time-delay system is more suitable for conducting the comparative analysis with similar dynamic time series, for instance, rainfall-runoff time series with time-delay responses. The remainder of the study is organized as follows. Section 2 presents the framework of the proposed model, the recurrent mechanism and the evolutionary algorithm. Section 3 shows detailed model evaluation on the benchmark time series. Section 4 presents the results and discussion on the proposed model applied to flood inflow forecasting. Conclusions are then drawn in Section 5.

2. Multi-step-ahead flood forecast methodology

2.1. Framework

The main goal of this study aims at exploring an evolutionary algorithm (GL) that hybrids the GA and the LSE to configure the recurrent ANFIS model into future forecast horizons for alleviating time-lag phenomena caused by the local search algorithm as well as improving forecast accuracy through derivative-free exploration with weight adjustment in training stages. Both the traditional ANFIS, denoted as ANFIS(SL), and the traditional recurrent ANFIS, denoted as R-ANFIS (SL), that combines the SDA and the LSE for model configuration are the comparative models in this study. Fig. 1 presents the architecture of the proposed R-ANFIS(GL) model, where the recurrent mechanism of the R-ANFIS illustrated in Fig. 1(a) is responsible for predicting future outcomes while the proposed evolutionary algorithm (GL) illustrated in Fig. 1(b) is responsible for searching the optimal parameters of the R-ANFIS. The construction of the recurrent mechanism and the evolutionary algorithm of the proposed R-ANFIS(GL) for modelling multi-step-ahead flood forecasts together with its overall effectiveness are presented as follows.

2.2. Recurrent mechanism

The recurrent mechanism in the R-ANFIS intends to forecast future outputs by using the inter feedback (forecasted value) of the model when the external feedback (observed value) is not available. The patterns between input and output variables of the ANFIS and the R-ANFIS are illustrated in Fig. 1(a) and described as follows.

For the ANFIS,

$$\text{Horizon t + 1: } \hat{Y}(t+1) = F(Y(t), Y(t-1), \cdots, Y(t-k), XX(t)) \quad (1a)$$

$$XX(t) = \{X(t), X(t-1), \cdots, X(t-d)\} \quad (1b)$$

$$\text{Horizon t + 2: } \hat{Y}(t+2) = F(Y(t), Y(t-1), \cdots, Y(t-k), XX(t)) \quad (2)$$

$$\text{Horizon t + 3: } \hat{Y}(t+3) = F(Y(t), Y(t-1), \cdots, Y(t-k), XX(t)) \quad (3)$$

...

$$\text{Horizon t + n: } \hat{Y}(t+n) = F(Y(t), Y(t-1), \cdots, Y(t-k), XX(t)) \quad (4)$$

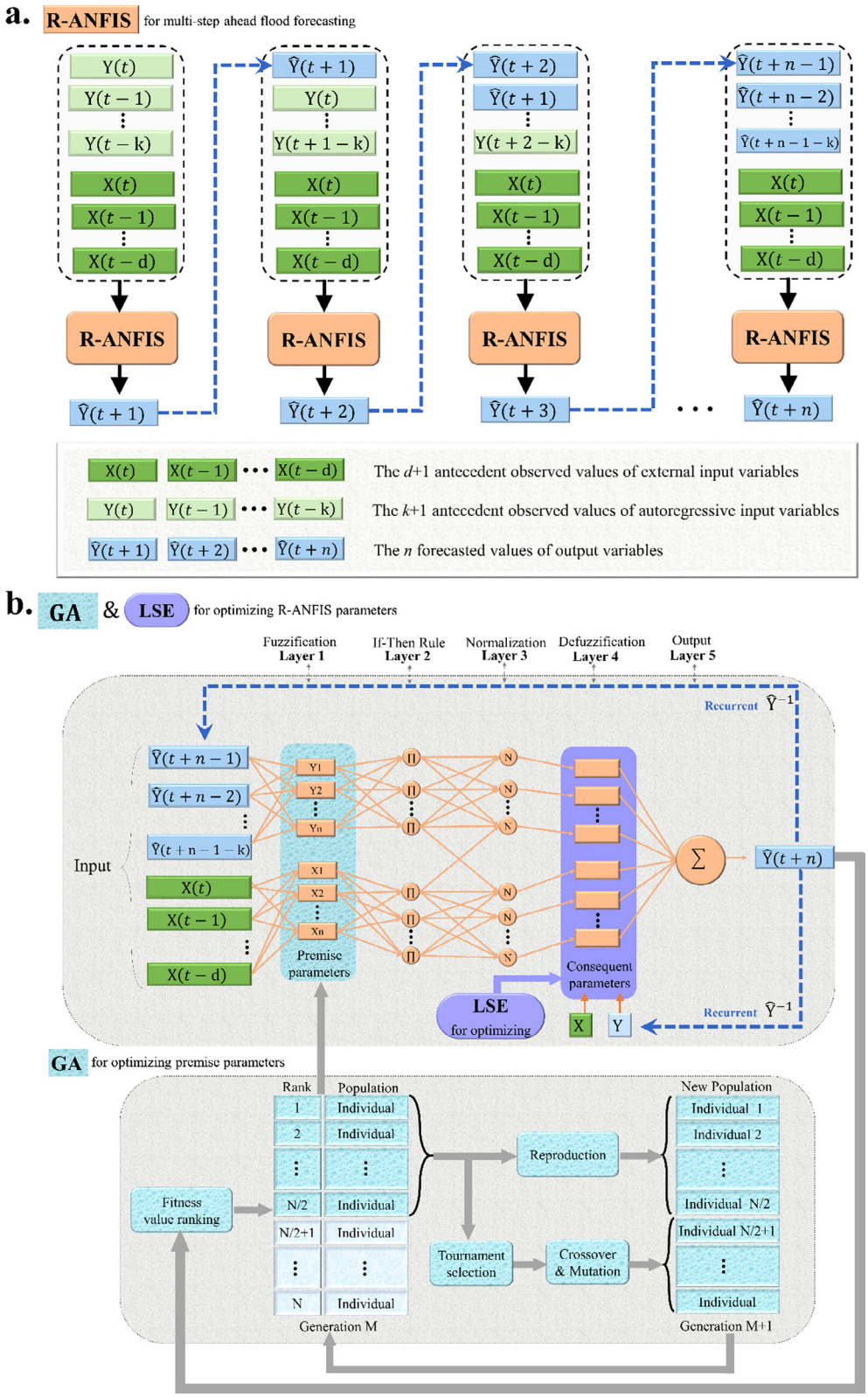

Fig. 1. R-ANFIS(GL) architecture. (a) R-ANFIS for multi-step-ahead forecasting. (b) Optimization of R-ANFIS parameters using the proposed evolutionary algorithm that fuses the Genetic Algorithm (GA) with the Least Square Estimator (LSE).

For the R-ANFIS,

$$\text{Horizon t + 1: } \hat{Y}(t+1) = F(Y(t), Y(t-1), \cdots, Y(t-k), XX(t)) \quad (5)$$

$$\text{Horizon t + 2: } \hat{Y}(t+2) = F(\hat{Y}(t+1), Y(t), \cdots, Y(t+1-k), XX(t)) \quad (6)$$

Horizon t + 3: $\hat{Y}(t+3) = F(\hat{Y}(t+2), \hat{Y}(t+1), \cdots, Y(t+2-k), XX(t))$ (7)

...

Horizon t + n: $\hat{Y}(t+n) = F(\hat{Y}(t+n-1), \cdots, \hat{Y}(t+2), \hat{Y}(t+n-1-k), XX(t))$ (8)

where $F(\cdot)$ is the function of the pattern between input and output variables. $\hat{Y}(t+1)$, $\hat{Y}(t+2)$, $\cdots, \hat{Y}(t+n)$ denote the forecasted values at horizons $t+1, t+2, \cdots, t+n$, respectively. $X(t)$, $X(t-1)$, $\cdots, X(t-d)$ denote the d + 1 antecedent observed values in external input variables.$XX(t)$ is the set of antecedent observed values in external input variables. $Y(t)$, $Y(t-1)$, $\cdots, Y(t-k)$ denote the k + 1 antecedent observed values in autoregressive input variables.

2.3. Evolutionary algorithm

The evolutionary algorithm (GL) proposed in this study for the R-ANFIS aims at optimizing the non-linear (premise) parameters in Layer 1 by using the GA and the linear (consequent) parameters in Layer 4 by using the LSE. The computation steps are illustrated in Fig. 1(b) and introduced as follows.

(1) Initialize population with real coding:

Randomly generate an initial population (Pop) for the non-linear parameters in Layer 1 (fuzzification). Each node in Layer 1 implements fuzzification by setting the membership degree of each input variable. Mathematically, the output of each node here can be expressed by Eqs. (9) and (10).

$$O_{1,i} = \mu_{A_i}(x_1), \quad \text{for } i = 1, 2, \quad \text{or} \tag{9}$$

$$O_{1,i} = \mu_{B_i}(x_2), \quad \text{for } i = 1, 2 \tag{10}$$

where x_1 and x_2 are the input variables. $O_{1,i}$ is the output of the ith node. $\mu_{A_i}(\cdot)$ and $\mu_{B_i}(\cdot)$ are the membership functions for fuzzy rules A_i and B_i, respectively.

The fuzzy set associated with each input node is characterized by the shape of its membership function. Membership functions can be any continuous and piecewise differentiable functions such as the Gaussian function as well as the generalized bell shaped, the triangular shaped and the trapezoidal shaped functions. Owing to smoothness and concise notation, the Gaussian and the generalized bell-shaped membership functions are increasingly popular for specifying fuzzy sets (Chang and Chang, 2006). A number of researches demonstrated that these two membership functions were more suitable for climatic and hydrological forecasting (Keskin et al. 2006; Zounemat-Kermani and Teshnehlab 2008; Wang et al. 2009; Talebizadeh and Moridnejad 2011; Goyal et al. 2014). It is noted that the Gaussian membership function is superior to the generalized bell shaped one in this study in terms of forecast accuracy by means of intensive trial-and-error procedures. In consequence, the membership function of the ANFIS employees the Gaussian function in this study, as expressed in Eq. (11).

$$\mu_{A_i}(x_1) = \exp\left(\frac{(x-c)^2}{2\sigma^2}\right) \tag{11}$$

where $\{\sigma, c\}$ is the set of nonlinear parameters in Layer 1. When there are N_1 input variables and N_2 membership functions in Layer 1, the number of non-linear parameters becomes N_3 ($=2 \times N_1 \times N_2$) and the number of fuzzy rules becomes $N_4 (=(N_2)^{N_1})$.

(2) Propagate forward from Layer 1 to Layer 4:

Each node in Layer 2 (if-then rule operation) carries out fuzzy-AND operation with T-norm operators (Jang 1993). The output of Layer 2 can be expressed by Eq. (12).

$$O_{2,i} = w_i = \mu_{A_i}(x_1) \times \mu_{B_i}(x_2), \quad \text{for } i = 1, 2 \tag{12}$$

where $O_{2,i}$ (or w_i) is the output of the ith node in Layer 2.

The output of the ith node in Layer 3 (normalization) is the ratio of the output of the ith node in Layer 2 to the sum of all outputs in Layer 2, shown as follows.

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \tag{13}$$

where $O_{3,i}$ (or $\bar{w}_i$) is the output of the ith node in Layer 3.

Each node in Layer 4 performs defuzzification with a linear function.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x_1 + q_i x_2 + r_i) \tag{14}$$

where $O_{4,i}$ is the output of the ith node in Layer 4. $\{p_i, q_i, r_i\}$ is the set of linear parameters of Layer 4. When there are N_5 output variables, the number of linear parameters becomes N_6 ($=N_2 \times (N_1 + N_5)$) in this layer.

(3) Optimize the set of linear parameters in Layer 4 using the LSE:

Given the values of non-linear parameters in Layer 1, Eq. (14) can be transformed into a matrix equation:

$$O_4 = PZ \tag{15}$$

where O_4 is the output vector of Layer 4. P is the output vector, and Z is the matrix of linear parameters of Layer 3. The LSE of Z, Z^*, seeks to minimize the squared error $\|P\,Z - O_4\|^2$ that forms the grounds for linear regression. The formula for Z^* uses the pseudo-inverse of Z.

$$Z^* = (P^T P)^{-1} P^T O_4 \tag{16}$$

where P^T is the transpose of P, and $(P^T P)^{-1} P^T$ is the pseudo-inverse of P. Z^* is the set of the optimal linear parameters in Layer 4, where the optimization of parameters is carried out by the LSE.

(4) Propagate forward to Layer 5 and evaluate population to save the best individual:

Given the optimal linear parameters, population is evaluated by calculating the error function described in Eq. (17).

$$f(s) = \frac{1}{2}\sum_{i=1}^{N} [e(i)]^2 = \frac{1}{2}\sum_{i=1}^{N} (Y_f(i) - Y_o(i))^2 \tag{17}$$

where $f(p)$ is the error function corresponding to the population of non-linear parameters; e_i is the residual error corresponding to the ith data; N is the total number of observed data; and $Y_f(i)$ and $Y_o(i)$ are the forecasts and observations of the ith data, respectively;

(5) Implement the genetic operation routine:

The process involves: (a) a tentative new population is duplicated from parent chromosomes through the reproduction operation. A higher fitness (i.e. the elitism preservation strategy) is preferentially chosen for survival by the tournament selection operation (Goldberg 1989; Goldberg and Deb 1991); (b) two parent chromosomes are recombined into new offspring chromosomes utilizing the crossover procedure with probability (P_c); and (c) a mutation operation with probability (P_m) can be performed for maintaining genetic diversity in the next generation. As compared with the deterministic sampling for individual selection, the tournament selection operation has the main merits that it is a stochastic sampling for individual selection and its elitism preservation mechanism can overcome the premature convergence bottleneck by preferentially choosing an individual with a higher fitness value. As a result, the tournament selection operation is implemented in this study for producing a new offspring population to boost the convergence of parameter optimization for the hydrological

model.

(6) Terminate the computation process subject to the stopping criteria:

Evaluate the population according to Steps 2–4. If the number of generation is less than the maximal generation (G_{max}), then perform Steps 2–6 again. Otherwise, terminate the computation process and deliver the optimal results.

3. Model evaluation by benchmark time series

To test and verify the performance of the proposed R-ANFIS(GL) on multi-step-ahead forecasts, the traditional ANFIS(SL) and the R-ANFIS (SL) are selected as benchmarks on the grounds that: (1) they have the same model structure (i.e., the ANFIS); (2) they have the same model parameters (the SDA and the LSE) to optimize; (3) the comparison between the ANFIS(SL) and the R-ANFIS(SL) is conducted at first to demonstrate the contribution of the recurrent mechanism to multi-step-ahead forecasting while the comparison between the R-ANFIS(SL) and the R-ANFIS(GA) is then implemented for demonstrating the contribution of the evolutionary algorithm (the GA) to multi-step-ahead flood forecasting. We intend to show that both the recurrent mechanism and the evolutionary algorithm play a pivotal role in ameliorating multi-step-ahead flood forecasting.

Besides, the famous Mackey-Glass time series utilized frequently to analyze the generalizability of different data-driven models or algorithms is selected as the benchmark time series in this study (Jang 1993; Ardalani-Farsa and Zolfaghari, 2010; Chang et al., 2012).

The Mackey-Glass series is expressed as follows.

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1 \times (t) \tag{18}$$

where the initial conditions are set as $x(0) = 1.2$, and $\tau = 17$.

In this study, a time series of x with variable length equal to 1000 is created by Eq. (18). During model construction, the leading 500 samples are utilized in the training stage, and the remaining 500 samples are divided equally into validation and testing stages, respectively. The numbers of variables for the embedding dimension (D) and time delay (T) in the Mackey-Glass time series are set as 4 and 6, respectively. We introduce a model input selection strategy in this study. When model training completes, three potential ANFIS networks with different structures are identified. Then the trained ANFIS network that produces the best performance in the validation stage is selected as the final model to be further tested by the testing dataset for evaluating model reliability. This model selection strategy is applied to the ANFIS(SL), the R-ANFIS(SL) and the R-ANFIS(GL), respectively.

The parameters of each ANFIS model in this study are summarized: (a) 4 ($=N_1$) input variables; (b) 2 ($=N_2$) membership functions; (c) 16 ($N_3 = 2 \times N_1 \times N_2$) non-linear parameters in Layer 1; (d) 16 ($N_4 = (N_2)^{N_1}$) fuzzy rules; (e) 1 ($=N_5$) output variable under the single-output pattern; and (f) 10 ($N_6 = N_2 \times (N_1 + N_5)$) linear parameters in Layer 4. In the SDA, the parameters of learning rate (η), decreasing factor (α), increasing factor (β) and maximum generation (G_{max}) are set as 0.01, 0.9, 1.1 and 1000, respectively. In the GA, the parameters of population (Pop), crossover probability (P_c), mutation probability (P_m) and maximum generation (G_{max}) are set as 1000, 0.9, 0.1 and 1000, respectively. It is manifest that the values of parameters required in the optimization algorithms could be obtained by means of intensive trial-and-error procedures for producing higher forecast accuracy.

Considering the stochastic nature of hydrological variables, one must not rely solely on any single indicator when evaluating the performance of hydrological forecast models. Additionally, for flood forecasting, it is very essential to know the performance of flow forecast models when forecasting high-magnitude data (e.g., reservoir inflows). Consequently, indicators utilized to evaluate model accuracy and forecast ability of flood peaks in this study consist of the Nash-Sutcliffe Efficiency coefficient (NSE, Nash 1970), the Root Mean Square Error (RMSE), the Goodness-of-Fit with respect to the benchmark (G_{bench}) and the Peak Percent Threshold Statistics (PPTS, Lohani et al., 2014). The formulae of the four indicators are expressed as follows.

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(Y_f(i) - Y_o(i))^2}{\sum_{i=1}^{N}(Y_o(i) - \bar{Y}_o)^2}, \quad \mathrm{NSE} \le 1 \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_f(i) - Y_o(i))^2}, \quad \mathrm{RMSE} \ge 0 \tag{20}$$

$$G_{bench} = 1 - \frac{\sum_{i=1}^{N}(Y_f(i) - Y_o(i))^2}{\sum_{i=1}^{N}(Y_o(i) - Y_{bench}(i))^2}, \quad G_{bench} \le 1 \tag{21}$$

$$\mathrm{PPTS}_{(l,u)} = \frac{1}{(k_l - k_u + 1)}\sum_{i=1}^{N}\left(\left|\frac{Y_f(i) - Y_o(i)}{Y_o(i)}\right|\right), \quad \mathrm{PPTS} \ge 0 \tag{22a}$$

$$k_l = \frac{l \times N}{100} \tag{22b}$$

$$k_u = \frac{u \times N}{100} \tag{22c}$$

where N is the number of the observed data, $\bar{Y}_o$ is the average of the observed data. $Y_{bench}(i)$ is the observed data shifted backwards by one or more time lags, for instance, $Y_{bench}(i) = Y_0(i - n)$ denotes the nth-step-ahead forecasted value. l and u are lower and higher limits in percentage, respectively.

It is clear from the definitions of these indicators that a model is considered to perform better if it has higher NSE and G_{bench} values while lower RMSE and PPTS values than the other comparative model (s).

Considering the time delay in the Mackey-Glass time series is 6 antecedent time steps ($=T$), the eight-step-ahead forecast (horizons t + 1 up to t + 8, time step = 1 h) is investigated for demonstrating the performance of the evolutionary R-ANFIS model (i.e. R-ANFIS(GL)). For the Mackey-Glass time series, the forecast results at different horizons are summarized in Table 1 and the residual errors of each model in validation and testing stages are illustrated in Fig. 2. The results of multi-step-ahead forecasts clearly show that:

1) The ANFIS(SL) performs the worst according to its comparatively larger RMSE & PPTS values but smaller G_{bench} & NSE values at different forecast horizons (Table 1). Besides, the band of its residual errors is the widest (Fig. 2). This is because the ANFIS(SL) does not utilize any recurrent mechanism and/or evolutionary algorithm (e.g., GL) such that the forecast error taking place in the initial timing of each recursion will cumulate and propagate to future recursions, which results in poor forecast accuracy.
2) The R-ANFIS(SL) and the R-ANFIS(GL) perform better than the non-recurrent ANFIS(SL) in all the validation and testing cases (Table 1, Fig. 2). It reveals that the recurrent mechanism of the ANFIS can adequately identify the dynamic and non-stationary properties of time series and therefore improves forecast accuracy.
3) Both the R-ANFIS(SL) and the R-ANFIS(GL) perform well (in terms of model stability and forecast accuracy) in training stages (Table 1). Nevertheless, the R-ANFIS(SL) performs less well than the R-ANFIS (GL) in the validation and testing stages due to the instability caused by its inherent derivative operation.
4) The R-ANFIS(GL) performs preeminently in all the cases according to RMSE, PPTS, G_{bench} and NSE values (Table 1) and the residual errors (Fig. 2). The results show that the proposed evolutionary recurrent model (R-ANFIS(GL)) in consideration of the nearest anterior feedback can efficiently modify synaptic weights, and therefore reliable and accurate multi-step-ahead forecasts can be achieved.

Table 1
Model performance of multi-step-ahead forecasts with respect to the Mackey-Glass time-series.

Period	Model	Indicator	Horizon							
			t + 1	t + 2	t + 3	t + 4	t + 5	t + 6	t + 7	t + 8
Training	ANFIS(SL)	PPTS(2%)	0.0312	0.0335	0.0374	0.0415	0.0447	0.0530	0.0589	0.0625
		PPTS(5%)	0.0328	0.0352	0.0393	0.0436	0.0469	0.0557	0.0618	0.0656
		PPTS(10%)	0.0343	0.0369	0.0411	0.0457	0.0492	0.0583	0.0648	0.0688
		PPTS(20%)	0.0374	0.0402	0.0449	0.0498	0.0536	0.0636	0.0707	0.0750
		RMSE	0.0021	0.0046	0.0073	0.0120	0.0208	0.0316	0.0431	0.0527
		NSE	0.998	0.988	0.944	0.896	0.839	0.732	0.645	0.552
		G_{bench}	0.998	0.982	0.862	0.811	0.727	0.605	0.477	0.429
	R-ANFIS(SL)	PPTS(2%)	0.0312	0.0325	0.0363	0.0403	0.0434	0.0515	0.0572	0.0607
		PPTS(5%)	0.0328	0.0342	0.0381	0.0423	0.0456	0.0540	0.0600	0.0637
		PPTS(10%)	0.0343	0.0368	0.0411	0.0456	0.0491	0.0582	0.0647	0.0686
		PPTS(20%)	0.0374	0.0401	0.0448	0.0497	0.0535	0.0635	0.0705	0.0749
		RMSE	0.0021	0.0042	0.0047	0.0062	0.0186	0.0214	0.0307	0.0392
		NSE	0.998	0.996	0.953	0.926	0.910	0.882	0.860	0.802
		G_{bench}	0.998	0.995	0.931	0.912	0.894	0.855	0.824	0.750
	R-ANFIS(GL)	PPTS(2%)	0.0294	0.032	0.034	0.036	0.037	0.043	0.047	0.049
		PPTS(5%)	0.0315	0.034	0.036	0.040	0.041	0.047	0.052	0.055
		PPTS(10%)	0.0342	0.0367	0.0410	0.0455	0.0490	0.0581	0.0646	0.0685
		PPTS(20%)	0.0373	0.0401	0.0447	0.0497	0.0535	0.0634	0.0705	0.0748
		RMSE	0.0018	0.0039	0.0042	0.0050	0.0122	0.0189	0.0255	0.0317
		NSE	0.998	0.997	0.991	0.982	0.961	0.941	0.923	0.907
		G_{bench}	0.998	0.997	0.988	0.962	0.944	0.931	0.917	0.901
Validation	ANFIS(SL)	PPTS(2%)	0.0314	0.0338	0.0377	0.0418	0.0451	0.0534	0.0594	0.0630
		PPTS(5%)	0.0330	0.0355	0.0396	0.0439	0.0473	0.0561	0.0623	0.0662
		PPTS(10%)	0.0346	0.0371	0.0415	0.0460	0.0496	0.0588	0.0653	0.0693
		PPTS(20%)	0.0377	0.0405	0.0452	0.0502	0.0541	0.0641	0.0712	0.0756
		RMSE	0.0021	0.0046	0.0073	0.0120	0.0210	0.0319	0.0435	0.0530
		NSE	0.998	0.986	0.941	0.893	0.831	0.729	0.640	0.547
		G_{bench}	0.998	0.982	0.862	0.811	0.725	0.595	0.472	0.425
	R-ANFIS(SL)	PPTS(2%)	0.0314	0.0328	0.0366	0.0406	0.0437	0.0518	0.0576	0.0611
		PPTS(5%)	0.0330	0.0344	0.0384	0.0426	0.0459	0.0544	0.0605	0.0642
		PPTS(10%)	0.0346	0.0370	0.0413	0.0459	0.0494	0.0586	0.0651	0.0691
		PPTS(20%)	0.0377	0.0404	0.0451	0.0500	0.0539	0.0639	0.0710	0.0754
		RMSE	0.0021	0.0042	0.0047	0.0062	0.0191	0.0217	0.0311	0.0397
		NSE	0.998	0.996	0.948	0.924	0.893	0.861	0.835	0.808
		G_{bench}	0.998	0.995	0.931	0.912	0.889	0.852	0.813	0.746
	R-ANFIS(GL)	PPTS(2%)	0.0296	0.0318	0.0342	0.0363	0.0374	0.0433	0.0474	0.0495
		PPTS(5%)	0.0317	0.0340	0.0365	0.0398	0.0410	0.0474	0.0518	0.0550
		PPTS(10%)	0.0344	0.0369	0.0412	0.0457	0.0493	0.0584	0.0649	0.0689
		PPTS(20%)	0.0375	0.0403	0.0450	0.0499	0.0537	0.0637	0.0708	0.0751
		RMSE	0.0018	0.0039	0.0042	0.0050	0.0125	0.0192	0.0258	0.0322
		NSE	0.998	0.997	0.991	0.974	0.950	0.937	0.922	0.907
		G_{bench}	0.998	0.997	0.988	0.962	0.943	0.926	0.912	0.893
Testing	ANFIS(SL)	PPTS(2%)	0.0315	0.0338	0.0378	0.0419	0.0451	0.0535	0.0595	0.0631
		PPTS(5%)	0.0331	0.0355	0.0397	0.0440	0.0474	0.0562	0.0625	0.0663
		PPTS(10%)	0.0347	0.0372	0.0416	0.0461	0.0497	0.0589	0.0654	0.0694
		PPTS(20%)	0.0378	0.0406	0.0453	0.0503	0.0542	0.0642	0.0714	0.0758
		RMSE	0.0025	0.0051	0.0075	0.0127	0.0213	0.0322	0.0448	0.0547
		NSE	0.996	0.981	0.936	0.871	0.825	0.716	0.632	0.540
		G_{bench}	0.996	0.979	0.860	0.806	0.722	0.591	0.467	0.411
	R-ANFIS(SL)	PPTS(2%)	0.0315	0.0329	0.0367	0.0407	0.0439	0.0520	0.0578	0.0613
		PPTS(5%)	0.0331	0.0345	0.0385	0.0428	0.0461	0.0546	0.0607	0.0644
		PPTS(10%)	0.0347	0.0372	0.0415	0.0461	0.0496	0.0588	0.0654	0.0694
		PPTS(20%)	0.0378	0.0406	0.0453	0.0502	0.0541	0.0642	0.0713	0.0757
		RMSE	0.0025	0.0045	0.0048	0.0065	0.0195	0.0220	0.0317	0.0408
		NSE	0.996	0.995	0.939	0.920	0.886	0.853	0.828	0.801
		G_{bench}	0.996	0.992	0.928	0.906	0.884	0.845	0.811	0.741
	R-ANFIS(GL)	PPTS(2%)	0.0297	0.0319	0.0343	0.0364	0.0376	0.0435	0.0476	0.0497
		PPTS(5%)	0.0318	0.0342	0.0367	0.0400	0.0412	0.0476	0.0521	0.0552
		PPTS(10%)	0.0339	0.0363	0.0404	0.0448	0.0482	0.0571	0.0634	0.0672
		PPTS(20%)	0.0373	0.0400	0.0446	0.0489	0.0527	0.0623	0.0692	0.0734
		RMSE	0.0021	0.0042	0.0045	0.0055	0.0127	0.0199	0.0261	0.0324
		NSE	0.997	0.995	0.991	0.970	0.950	0.932	0.918	0.902
		G_{bench}	0.997	0.994	0.985	0.959	0.942	0.929	0.907	0.890

5) The results indicate that the RMSE and the PPTS gradually increase (the G_{bench} and the NSE gradually decrease) as the forecast horizon increases (i.e. from t + 1 to t + 8), and the largest increments of the RMSE and the PPTS along the forecast horizon for the three models are produced by the ANFIS(SL) while the smallest ones are produced by the R-ANFIS(GL).

To show the merit (recurrent mechanism and evolutionary approach) of the proposed R-ANFIS(GL) by taking the horizon of t + 8 for example, the results of the three ANFIS models in training stages are assessed (Fig. 3). It indicates that the error function values of the R-ANFIS(SL) and the R-ANFIS(GL) are smaller than those of the ANFIS (SL), where sharp drops occur around the 50th generation for the two

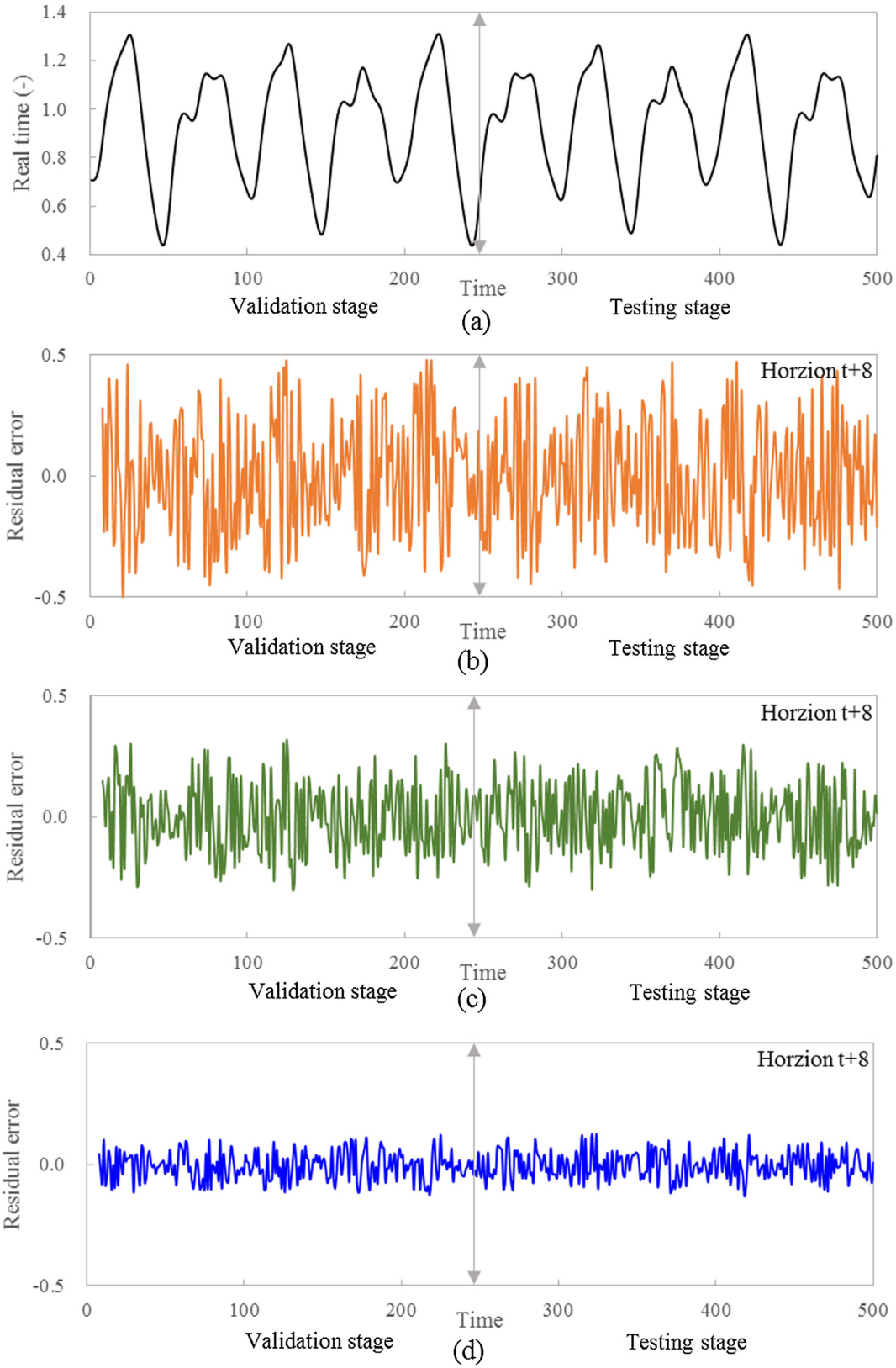

Fig. 2. (a) Mackey-Glass time-series. (b)-(d) Residual errors of the ANFIS(SL), R-ANFIS(SL) and R-ANFIS(GL) in validation and testing stages at horizon t + 8, respectively.

recurrent models. This is because the recurrent mechanism utilizes anterior observations and/or forecasted values to reduce error accumulation and propagation. It is noted that the error function values of the R-ANFIS(SL) and the R-ANFIS(GL) show less fluctuation than those of the ANFIS(SL). The error function of the R-ANFIS(GL) even produces a smoother curve with a monotonic decreasing trend, which implies the ANFIS(GL) is able to avoid falling into the trap of local minima. In addition, the initial error function value in the R-ANFIS(GL) is

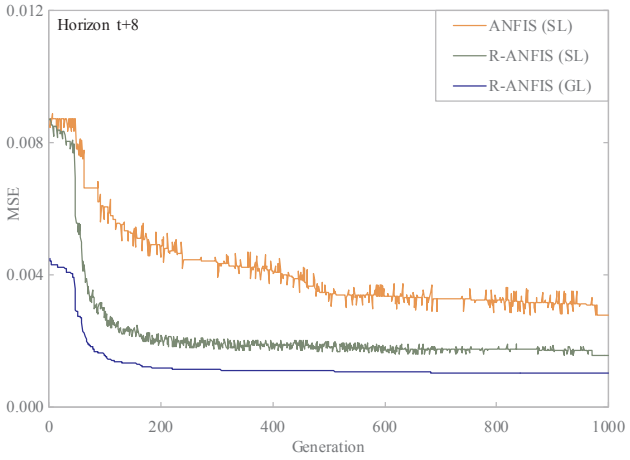

Fig. 3. Error function values of ANFIS models in training stage at horizon t + 8 (Error function is the Mean Squared Error with normalized dataset).

significantly smaller than those of the other two models. This is resulting from the evolutionary approach that the elitism preservation strategy of the GL algorithm could evaluate and select the optimal solution (ANFIS parameters) from 1000 solutions (=Pop) in every generation. The SL algorithm, on the contrary, could only move forward one by one solution in every generation. These results clearly show that the proposed evolutionary algorithm (i.e., GL) can conquer the shortcomings of model instability and local minima caused by the local search algorithm (i.e. SL).

We would like to note that the main differences between the local search algorithm and the proposed evolutionary algorithm are: (1) the local search algorithm combines the SDA with the LSE (Jang 1993; Zhang and Morris 1999) while the proposed evolutionary algorithm combines the GA with the LSE; (2) the local search algorithm is a derivative and deterministic optimization algorithm (Tamura et al. 2008) while the proposed evolutionary algorithm is a derivative-free and stochastic optimization algorithm; and (3) the local search algorithm is a non-evolutionary algorithm confined by the initialization of the non-linear parameters of the ANFIS while the proposed evolutionary algorithm is an AI-based global search algorithm.

4. Application

We next evaluate the applicability and reliability of the proposed ANFIS(GL) model with a reservoir inflow series. A brief introduction of the study area, illustrated in Fig. 4, is made as follows. The Chang-Jiang River extending 6300 km is the longest river in China, and its drainage area covers 1.80 million km^2. The Chang-Jiang River is located in the subtropical zone between the Indian Ocean and the North Pacific Ocean. Floods in this basin usually result from heavy rainfalls such that downstream flooding may occur just within one day. Therefore, reservoir operation targeting river-basin flood control and water resource management demands for accurate multi-step-ahead flood forecasts that can adequately deal with the high variability of river flow. In the Chang-Jiang River Basin, the Xiang-Jia-Ba (XJB) Reservoir in the upstream and the Three Gorges Reservoir (TGR) constitute a top-down cascade. These two reservoirs are pivotal hydraulic facilities in this basin and serve multiple purposes of flood control, hydropower generation, navigation, etc. The XJB and the TGR (the largest reservoir in the world to date) have drainage areas of 0.46 and 1.00 million km^2 accordingly, and their flood control capacities are of 0.90 and 22.15 billion m^3, respectively. Because of pervasive human activities across the earth and global warming, it is difficult to find a watershed with hydrological systems unaffected by a diversity of natural and/or human powers, for instance, climatic variability and change, water resources engineering projects and land cover/use change (Jiang et al., 2015; Gao et al., 2017; Shen et al., 2018). According to the researches of the Three Gorges Reservoir (TGR) (Bai et al., 2016; Jiang et al., 2017), rainfall and runoff variables were found non-stationary due to climate change and human activities. Therefore, it is expected that novel and sophisticated data-driven models can be developed to model the notoriously non-linear time series in this study.

Fig. 4 shows the distribution of reservoirs, rivers, streamflow gauge stations and rainfall gauge stations in the study area. The reservoir inflow and rainfall data are available at a time step of 6 h collected from 2003 to 2016. The observed data of 67 rainfall gauge stations in Regions I (13 stations colored in red in Fig. 4) and II (54 stations colored in blue in Fig. 4) are utilized for calculating the weighted average areal rainfall in these two regions separately. A total of 40,992 data (12 (input variables) × 7 years (2003–2009) × 488 data/year (June 1st to September 30th)) are adopted for training the models whereas 17,568 data (12 (input variables) × 3 years (2010–2012) × 488 data/year (June 1st to September 30th)) and 23,424 data (12 (input variables) × 4 years (2013–2016) × 488 data/year (June 1st to September 30th)) are adopted for validating and testing the models, respectively.

Because the travel distance and the travel time of flow running from each upstream station to the TGR differs, a statistical indicator is needed to identify the relationship between the flow (or rainfall) of each upstream station and the inflow of the TGR. In comparison with other model input selection methods (e.g., the Pearson coefficient, the Spearman coefficient, and the principal component analysis), the Kendall tau coefficient (Maidment, 1993) has the advantages that: (1) the investigative data do not need to satisfy the hypothesis relevant to the requirement for data to be normally distributed; (2) it is commonly used to analyze the characteristics of the non-linear correlation between two datasets; and (3) it has wider applicability owing to its ability of non-parametric statistical analysis (Méheust et al., 2012; Chen et al., 2013; Acharya et al., 2014; Lebrenz and Bárdossy, 2017; Paudel et al., 2017). Therefore, the Kendall tau coefficient analysis is conducted in this study to identify the highest correlation regarding the time lags between input and output variables. According to the highest Kendall tau rank correlation coefficients, the time lags between the inflow of the TGR and flow gauge stations as well as areal rainfall are set as 48 h (XJB reservoir), 48 h (F_1), 48 h (F_2), 42 h (F_3), 42 h (F_4), 24 h (F_5), 18 h (F_6), 18 h (F_7), 12 h (F_8), 42 h (Rainfall-I) and 12 h (Rainfal-II), respectively. The parameters in each ANFIS model are summarized: (a) 12 (=N_1) input variables; (b) 2 (=N_2) membership functions; (c) 48 ($N_3 = 2 \times N_1 \times N_2$) non-linear parameters in Layer 1; (d) 4096 ($N_4 = (N_2)^{N_1}$) fuzzy rules; (e) 1 (=N_5) output variable under the single-output pattern; and (f) 26 ($N_6 = N_2 \times (N_1 + N_5)$) linear parameters in Layer 4. In the SDA, the parameters of learning rate (η), decreasing factor (α), increasing factor (β) and maximum generation (G_{max}) are set as 0.01, 0.9, 1.1 and 1000, respectively. In the GA, the parameters of population (Pop), crossover probability (P_c), mutation probability (P_m) and maximum generation (G_{max}) are set as 1000, 0.9, 0.1 and 1000, respectively.

Considering the travel time of flow running from each upstream station to the TGR varies between 6 h and 48 h, the model performance of one- up to eight-step-ahead forecasting (horizons t + 1 up to t + 8, time step = 6 h) are investigated. Moreover, the PPTS can project model performance onto various magnitude ranges of the data. PPTS (l,100) indicates the peak percent threshold statistics of top l% data. Therefore, the PPTS is used to verify model performance for the high-magnitude data in this study. For this purpose, PPTS values associated with the highest 2%, 5%, 10% and 20% data have been calculated, respectively. Table 2 summarizes forecast results of the three investigative models. It appears that the R-ANFIS(GL) model produces much higher NSE & G_{bench} values but much smaller RMSE & PPTS values than the ANFIS(SL) model in all the three (training, validation and testing) stages. It is noted that for the ANFIS(SL) model, the RMSE & the PPTS significantly increase and the NSE & the G_{bench} effectively

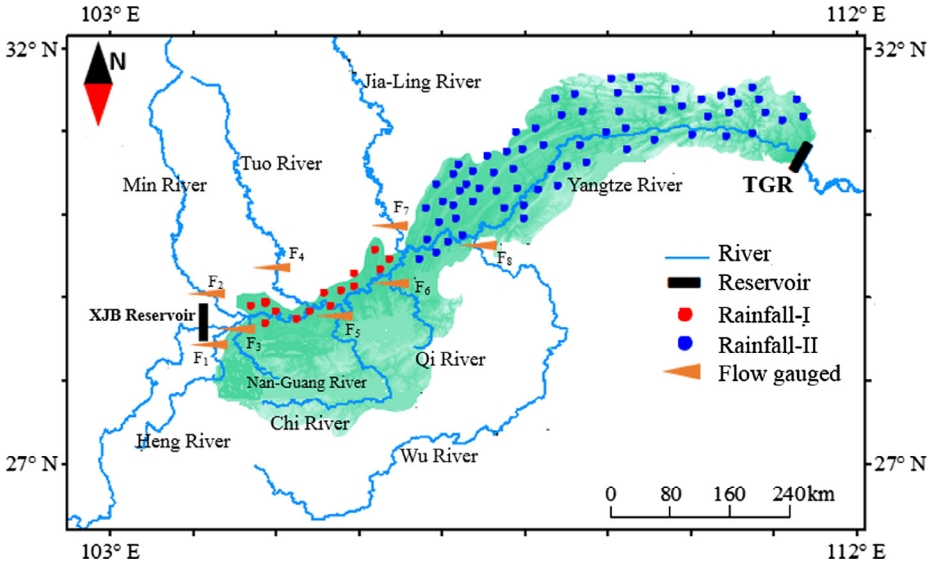

Fig. 4. Location of the Three Gorges Reservoir (TGR), flow and rainfall gauge stations.

Table 2
Model performance of multi-step-ahead forecasts at with respect to reservoir inflow (ANFIS(SL) vs. R-ANFIS(GL)).

Period	Model	Indicator	Horizon							
			t + 1	t + 2	t + 3	t + 4	t + 5	t + 6	t + 7	t + 8
Training	ANFIS(SL)	PPTS(2%)	0.0321	0.0348	0.0393	0.0423	0.0478	0.0562	0.0636	0.0681
		PPTS(5%)	0.0337	0.0366	0.0412	0.0444	0.0502	0.0590	0.0668	0.0715
		PPTS(10%)	0.0353	0.0383	0.0432	0.0466	0.0526	0.0618	0.0700	0.0749
		PPTS(20%)	0.0386	0.0418	0.0471	0.0508	0.0574	0.0674	0.0763	0.0818
		RMSE	867	1063	1218	1472	1862	2026	2301	2452
		NSE	0.964	0.955	0.936	0.897	0.859	0.794	0.716	0.659
		G_{bench}	0.909	0.869	0.819	0.749	0.689	0.619	0.539	0.460
	R-ANFIS(GL)	PPTS(2%)	0.0306	0.0315	0.0349	0.0399	0.0412	0.0489	0.0538	0.0546
		PPTS(5%)	0.0321	0.0331	0.0366	0.0419	0.0433	0.0513	0.0564	0.0573
		PPTS(10%)	0.0336	0.0357	0.0394	0.0451	0.0466	0.0553	0.0608	0.0618
		PPTS(20%)	0.0367	0.0389	0.0430	0.0492	0.0509	0.0603	0.0663	0.0674
		RMSE	746	909	1074	1241	1412	1584	1741	1838
		NSE	0.983	0.972	0.966	0.945	0.928	0.905	0.886	0.823
		G_{bench}	0.931	0.921	0.881	0.851	0.812	0.772	0.723	0.663
Validation	ANFIS(SL)	PPTS(2%)	0.0324	0.0355	0.0400	0.0439	0.0469	0.0572	0.0641	0.0693
		PPTS(5%)	0.0340	0.0372	0.0420	0.0461	0.0492	0.0600	0.0673	0.0728
		PPTS(10%)	0.0356	0.0390	0.0440	0.0483	0.0515	0.0629	0.0705	0.0762
		PPTS(20%)	0.0389	0.0425	0.0480	0.0527	0.0562	0.0686	0.0769	0.0832
		RMSE	890	1081	1350	1505	1928	2041	2365	2555
		NSE	0.963	0.954	0.935	0.896	0.858	0.793	0.715	0.658
		G_{bench}	0.899	0.849	0.789	0.739	0.669	0.599	0.529	0.440
	R-ANFIS(GL)	PPTS(2%)	0.0296	0.0317	0.0341	0.0362	0.0374	0.0433	0.0473	0.0494
		PPTS(5%)	0.0316	0.0340	0.0365	0.0398	0.0410	0.0473	0.0517	0.0549
		PPTS(10%)	0.0344	0.0369	0.0412	0.0457	0.0492	0.0584	0.0649	0.0688
		PPTS(20%)	0.0375	0.0402	0.0449	0.0498	0.0537	0.0637	0.0707	0.0751
		RMSE	763	928	1119	1285	1525	1603	1786	1932
		NSE	0.981	0.970	0.964	0.943	0.926	0.903	0.884	0.821
		G_{bench}	0.929	0.919	0.879	0.849	0.809	0.769	0.709	0.649
Testing	ANFIS(SL)	PPTS(2%)	0.0325	0.0355	0.0404	0.0448	0.0479	0.0573	0.0642	0.0694
		PPTS(5%)	0.0341	0.0373	0.0424	0.0471	0.0502	0.0601	0.0675	0.0729
		PPTS(10%)	0.0357	0.0391	0.0445	0.0493	0.0526	0.0630	0.0707	0.0764
		PPTS(20%)	0.0389	0.0426	0.0485	0.0538	0.0574	0.0687	0.0771	0.0833
		RMSE	902	1097	1397	1595	1963	2096	2389	2684
		NSE	0.964	0.955	0.936	0.890	0.859	0.794	0.716	0.659
		G_{bench}	0.899	0.829	0.779	0.699	0.629	0.559	0.490	0.420
	R-ANFIS(GL)	PPTS(2%)	0.0294	0.0313	0.0333	0.0350	0.0357	0.0409	0.0443	0.0462
		PPTS(5%)	0.0315	0.0335	0.0356	0.0384	0.0392	0.0448	0.0484	0.0514
		PPTS(10%)	0.0335	0.0355	0.0396	0.0439	0.0468	0.0548	0.0608	0.0645
		PPTS(20%)	0.0373	0.0400	0.0446	0.0489	0.0527	0.0623	0.0692	0.0734
		RMSE	775	937	1188	1307	1563	1626	1825	2004
		NSE	0.983	0.972	0.966	0.945	0.928	0.905	0.886	0.823
		G_{bench}	0.929	0.919	0.869	0.839	0.799	0.749	0.689	0.619

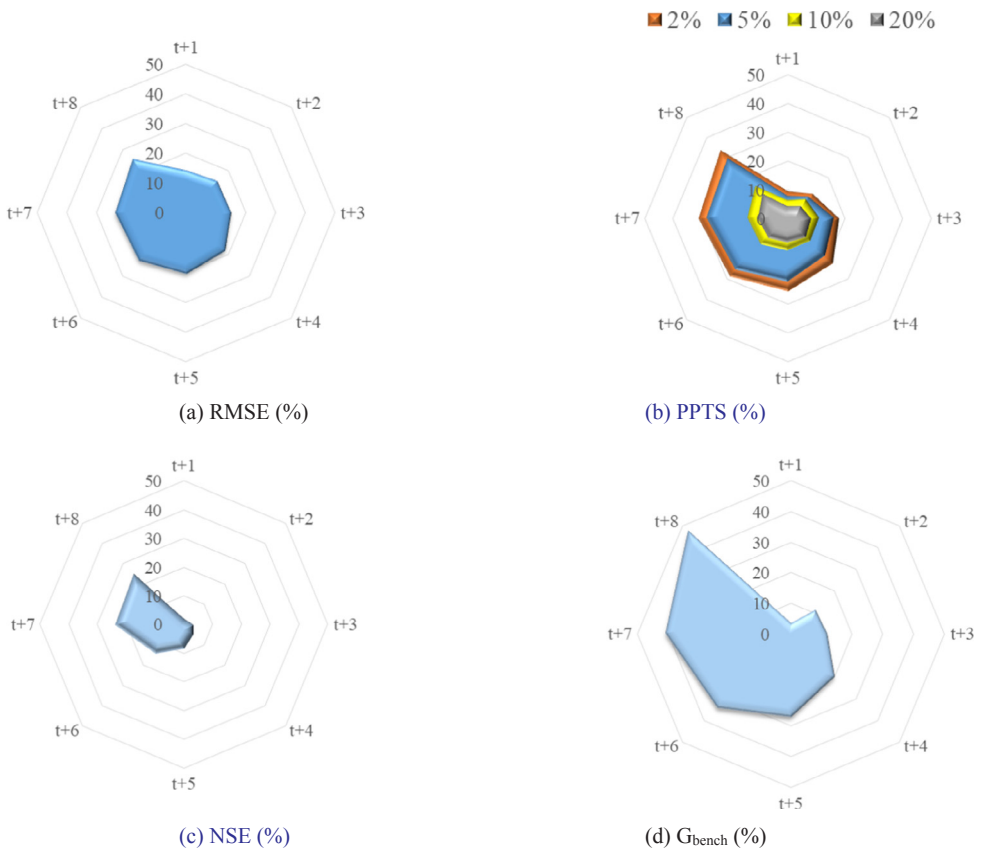

Fig. 5. Improvement rates in terms of RMSE, PPTS, NSE and G_{bench} in the testing stage of the R-ANFIS(GL) for multi-step-ahead flood forecasts at the Three Gorges Reservoir (TGR), as compared with the ANFIS(SL).$\text{Improvementrateofeachindicator} = \frac{|\text{Indicator(ANFIS(SL))} - \text{Indicator(R} - \text{ANFIS(GL))}|}{\text{Indicator(ANFIS(SL))}} \times 100\%$

decrease in all stages at earlier time steps (from t + 2 up to t + 5). In contrast, the values of the RMSE, the PPTS, the NSE and the G_{bench} for the R-ANFIS(GL) model remain relatively stable. The results show that the recurrent (delayed) outputs of the proposed model do provide valuable information to increase multi-step-ahead flood forecast accuracy. It is depicted further from Table 2 that the PPTS increases gradually as the forecast horizon increases from t + 1 to t + 8, and the largest increments in PPTS along the forecast horizon are made by the ANFIF (SL) while the smallest ones are made by the R-ANFIS(GL). In addition, the R-ANFIS(GL) illustrates its preeminence over the ANFIF (SL) in forecasting high values of flooding at 2% and 5% frequencies while both models show similar performance at 10% and 20% frequencies. It is noticed that different data sets would cause different variations in PPTS, which could provide an alternative solution to critically selecting a suitable flood forecasting model.

We then assess the model predictability for different forecast horizons. According to Table 2, the R-ANFIS(GL) model produces better performance in all the three (training, validation and testing) stages, whereas the ANFIS(SL) model performs well only in the training stage at horizons up to t + 3 (NSE is higher than 0.90, and G_{bench} is higher than 0.80). The results show that the proposed evolutionary algorithm (combines GA and LSE) for the R-ANFIS(GL) model makes impressive achievements in model stability and generalizability. Besides, the ANFIS(SL) model produces small NSE and G_{bench} values (lower than 0.90 and 0.65 at horizon t + 5 and even lower than 0.70 and 0.45 at horizon t + 8, respectively) in testing datasets, whereas the R-ANFIS (GL) model produces high NSE and G_{bench} values (higher than 0.94 and 0.80 at horizons up to t + 4 (one day ahead) and higher than 0.80 and 0.60 at horizon t + 8 (two days ahead), respectively) in testing datasets. Given the maximal lag time, i.e. horizon t + 8 (two-day ahead), the R-ANFIS(GL) model can improve the NSE and the G_{bench} by 24.89% and 47.62% as well as reduce the RMSE and the PPTS(2%) values by 25.34% and 33.43% in the testing stage, respectively, as compared to the ANFIS(SL). Apparently, the predictability of the R-ANFIS(GL) model for future forecast horizons is significantly better than that of the ANFIS (SL) model. As shown in Fig. 5, there is an interesting finding that the improvement rates in terms of RMSE, PPTS, NSE and G_{bench} values significantly increase at horizons from t + 1 to t + 8. In other words, the R-ANFIS(GL) can gain more advantages than the ANFIS(SL) as the forecast horizon increases. From the perspective of indicator function, the RMSE, the PPTS and the G_{bench} are sensitive to mid-high flow magnitudes while the NSE is sensitive to flood volume. That means the proposed R-ANFIS(GL) model not only could largely increase forecast accuracy at mid-high flow magnitudes but also could improve the goodness-of-fit to flood volume at the same time. The results demonstrate that the R-ANFIS(GL) model is able to produce more stable and accurate multi-step-ahead forecasts by means of extracting the features of the non-stationary processes between rainfall and runoff. These improvements made by the R-ANFIS(GL) provide substantial evidence and much more confidence in flood control and water resources management for the TGR because the two-day-ahead forecasting is extremely critical and valuable to the TGR.

To clearly differentiate the capability of the ANFIS(SL) and R-ANFIS (GL) models, a flood event with maximal peak-flow reaching 69,100 cms is selected to test both models through assessing the goodness-of-fit between observations and forecasts, as shown in Fig. 6. It indicates that the R-ANFIS(GL) is able to forecast well at horizons t + 5 up to t + 8, whereas the ANFIS(SL) model has obvious time-lag phenomena as well as larger gaps between observations and forecasts. That is to say, the ANFIS(SL) model fails to forecast inflow values adequately at horizons more than t + 5. It shows that the R-ANFIS(GL) model is able to effectively trace the trails of flood events, significantly mitigate time-lag effects, and produce much accurate and reliable multi-step-ahead flood forecasts. Nevertheless, all the uncertainties

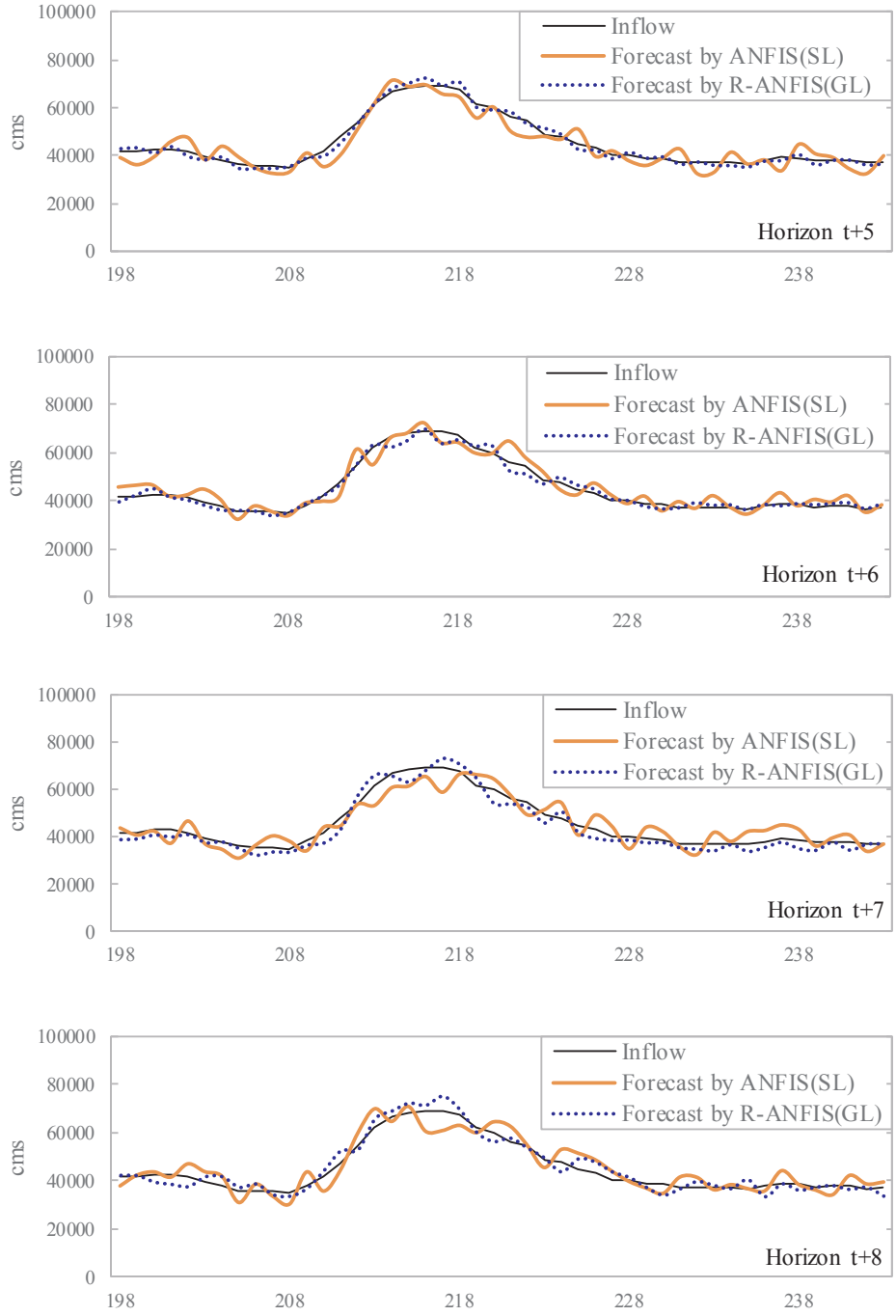

Fig. 6. Multi-step-ahead flood forecasts for the Three Gorges Reservoir (TGR) (a flood event with maximal peak-flow exceeding 69,000 cms was selected for testing the constructed models).

residing in model input, model structure and parameters could be the sources that cause the time-lag effects of the ANFIS(SL) and R-ANFIS (SL). In this study, we pay more attention to the exploration of the recurrent mechanism (i.e., model structure) and model parameter optimization (i.e., the evolutionary algorithm) for making multi-step-ahead flood forecasts through conquering the overfitting and instability bottlenecks. It is noted that the reasons for causing the time-lag effects of the ANFIS(SL) model consist of: (1) the instability phenomenon occurs at the ANFIS(SL) model; (2) for the Chang-Jiang River Basin with fast urban development, its regional meteorology frequently interacts with intensive human activities and climate change that gives rise to the non-stationary processes between rainfall and runoff while the ANFIS(SL) with a static learning mechanism is not able to well capture non-stationary features; and (3) for forecast horizons from t + 1 up to t + 8, no real-time information of rainfall or runoff is available to the ANFIS model while the R-ANFIS model is able to effectively utilize the forecasted runoffs from horizon t + 1 up to horizon t + 7 by dint of the recurrent mechanism that feeds the previous model outputs back to the input layer through a time-delay memory of k units (see Fig. 1). It is also noted that the proposed ANFIS(GL) model could alleviate the time-lag effects of multi-step-ahead forecasts but could not fully eliminate the effects. That is to say, to further improve forecast accuracy, it would need to fuse the forecasted values of external input variables (e.g., rainfall) into the multi-step-ahead flood forecast model.

5. Conclusions

In this study, we explored an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts. The demand for adaptation and evolutionary strategies for configuring ANFIS neural networks is due to real-world applications in modelling non-stationary environments. Three neuro-fuzzy learning models, including a traditional ANFIS that uses the steepest descent algorithm and the least square estimator (ANFIS(SL)), a recurrent ANFIS (R-ANFIS(SL)), and an R-ANFIS(GL) that uses an evolutionary optimization algorithm, are configured to model time series. These constructed models were assessed and compared using a famous benchmark chaotic time series as well as historical reservoir inflow series of the Three Gorges Reservoir in China. The main merit of the proposed R-ANFIS(GL)) neural network not only lies in the avoidance of the local minima problems and model instability but in reducing error accumulation and propagation faced in multi-step-ahead forecasting.

The results of the three models applied separately to the Mackey-Glass time series demonstrated that the R-ANFIS(GL) prominently outperformed the other two comparative models for all the training, validation and testing cases at different horizons (Table 1 and Fig. 2). The results also indicated that the increments of the RMSE along forecast horizon (i.e. from t + 1 to t + 8) of three models were the largest for the ANFIS(SL) while the smallest for the R-ANFIS(GL). Besides, the differences in RMSE values between the ANFIS(SL) and R-ANFIS(GL) increased with forecast horizon, similarly for G_{bench} values. It indicates that small prediction errors happening in the beginning would cumulate and propagate to the future when using the derivative optimization algorithm (i.e. SL) in the ANFIS, which would thus reduce the multi-step-ahead forecast accuracy.

According to the assessment on the two flood forecast models (i.e. the traditional ANFIS(SL) and our proposed R-ANFIS(GL) established for the TGR, the results (Table 2 and Fig. 5) clear indicate that the proposed R-ANFIS(GL) model could provide much better forecast the inflow series in the long forecast horizon and significantly mitigate time-lag phenomena than those of the ANFIS(SL). The reason that the traditional ANFIS(SL) model failed to achieve satisfactory multi-step-ahead forecast results in training, validation and testing stages could result from the instability caused by the inherent derivative operation. That is to say, the ANFIS demands for more sophisticated techniques, such as recurrent mechanism and/or evolutionary algorithm, to increase model stability and generalizability.

Because the ultimate goal concerns real-time forecast accuracy, the results of this study demonstrated that the proposed R-ANFIS(GL) model could provide reliable as well as precise multi-step-ahead flood forecasts owing to two key strategies: the incorporation of the antecedent observed values of external input variables into the recurrent mechanism for mitigating error accumulation and propagation; and the evolutionary operation of the GL for parameter optimization. In technical aspects, the output of the static neural network (ANFIS(SL)) relied solely on observed data whereas the output of the recurrent neural networks (R-ANFIS(GL)) would depend upon the optimal integration of observed data and forecasted data with time-delay units through feedback connections and thus significant contribution could be made to model outputs. The recurrent neuro-fuzzy networks have the merit of effectively extracting the input/output dependency and dynamic process owing to their recursive outputs and fuzzy clustering mechanism. From the perspectives that cause time lags, the recurrent neuro-fuzzy networks, however, could not fully eliminate the time-lag bottleneck encountered in multi-step-ahead forecasting due to the complex interrelation and lacking observed rainfall/runoff input values. Therefore, our future work will focus on the incorporation of the forecasted values of external input variables (e.g., rainfall) into multi-step-ahead flood forecasting for further enhancing forecast accuracy. Additionally, this study concentrates on deterministic data-driven models for multi-step-ahead flood forecasting while more state-of-the-art AI techniques, for instance, the ensemble and probabilistic forecast techniques, could be integrated into data-driven models to reduce the uncertainties in multi-step-ahead flood forecasting in future research.

Conflict of interest

None declared.

Acknowledgements

This study is financially supported by the China Postdoctoral Science Foundation (No. 2017M620336), the Research Council of Norway (FRINATEK Project 274310), the China National Key Research and Development Program (2018YFC0407904) and the Ministry of Science and Technology, Taiwan (MOST 104-2313-B-002 -023 -MY3). The authors would like to thank the editors and anonymous reviewers for their valuable and constructive comments related to this manuscript.

References

Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Wilby, R.L., 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. Prog. Phys. Geogr. 36 (4), 480–513.

Acharya, N., Shrivastava, N.A., Panigrahi, B.K., Mohanty, U.C., 2014. Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine. Clim. Dyn. 43 (5–6), 1303–1310.

Alexander, A.A., Thampi, S.G., 2018. Development of hybrid wavelet-ANN model for hourly flood stage forecasting. ISH J. Hydraul. Eng. 1–9.

Ardalani-Farsa, M., Zolfaghari, S., 2010. Chaotic time series prediction with residual analysis method using hybrid Elman-NARX neural networks. Neurocomputing 73 (13), 2540–2553.

Bai, Y., Chen, Z., Xie, J., Li, C., 2016. Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models. J. Hydrol. 532, 193–206.

Banihabib, M.E., Arabi, A., Salha, A.A., 2015. A dynamic artificial neural network for assessment of land-use change impact on warning lead-time of flood. Int. J. Hydrol. Sci. Technol. 5 (2), 163–178.

Chandra, R., 2015. Competition and collaboration in cooperative coevolution of Elman recurrent neural networks for time-series prediction. IEEE Trans. Neural Networks Learn. Syst. 26 (12), 3123–3136.

Chang, F.J., Chang, Y.T., 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Adv. Water Resour. 29 (1), 1–10.

Chang, L.C., Chang, F.J., Wang, K.W., Dai, S.Y., 2010. Constrained genetic algorithms for optimizing multi-use reservoir operation. J. Hydrol. 390 (2), 66–74.

Chang, L.C., Chen, P.A., Chang, F.J., 2012. Reinforced two-step-ahead weight adjustment technique for online training of recurrent neural networks. IEEE Trans. Neural Networks Learn. Syst. 23 (8), 1269–1278.

Chang, F.J., Tsai, M.J., 2016. A nonlinear spatio-temporal lumping of radar rainfall for modeling multi-step-ahead inflow forecasts by data-driven techniques. J. Hydrol. 535, 256–269.

Chang, F.J., Chen, P.A., Lu, Y.R., Huang, E., Chang, K.Y., 2014. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. J. Hydrol. 517, 836–846.

Chen, Y.H., Chang, F.J., 2009. Evolutionary artificial neural networks for hydrological systems forecasting. J. Hydrol. 367, 125–137.

Chen, P.A., Chang, L.C., Chang, F.J., 2013. Reinforced recurrent neural networks for multi-step-ahead flood forecasts. J. Hydrol. 497, 71–79.

Cheng, C.T., Wu, X.Y., Chau, K.W., 2005. Multiple criteria rainfall–runoff model calibration using a parallel genetic algorithm in a cluster of computers/Calage multi-critères en modélisation pluie–débit par un algorithme génétique parallèle mis en œuvre par une grappe d'ord inateurs. Hydrol. Sci. J. 50 (6), 1069–1087.

Ehteram, M., Mousavi, S.F., Karami, H., Farzin, S., Singh, V.P., Chau, K.W., El-Shafie, A., 2018. Reservoir operation based on evolutionary algorithms and multi-criteria decision-making under climate change and uncertainty. J. Hydroinf. 20 (2), 332–355.

Fazel, S.A.A., Mirfenderesk, H., Tomlinson, R., Blumenstein, M., 2015. Towards robust flood forecasts using neural networks. In: Neural Networks (IJCNN), 2015 International Joint Conference on. IEEE, pp. 1–6.

Fei, J., Lu, C., 2018. Adaptive sliding mode control of dynamic systems using double loop recurrent neural network structure. IEEE Trans. Neural Networks Learn. Syst. 29 (4), 1275–1286.

Firat, M., Güngör, M., 2008. Hydrological time-series modelling using an adaptive neuro-fuzzy inference system. Hydrol. Process. 22 (13), 2122–2132.

Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K.W., Faizollahzadeh Ardabili, S., Piran, M.J., 2018. Survey of computational intelligence as basis to big flood management: challenges, research directions and future work. Eng. Appl. Comput. Fluid Mech. 12 (1), 411–437.

Gao, S., Liu, P., Pan, Z., Ming, B., Guo, S., Xiong, L., 2017. Derivation of low flow frequency distributions under human activities and its implications. J. Hydrol. 549,

294–300.
Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning, second ed. Addison-Wesley, Reading, MA.
Goldberg, D.E., Deb, K., 1991. A comparative analysis of selection schemes used in genetic algorithms. In: Rawlins, G. (Ed.), Foundations of genetic algorithms. In Foundations of Genetic Algorithms. Morgan Kaufmann, San Mateo, CA, pp. 69–93.
Goyal, M.K., Bharti, B., Quilty, J., Adamowski, J., Pandey, A., 2014. Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. Expert Syst. Appl. 41 (11), 5267–5276.
He, Z., Wen, X., Liu, H., Du, J., 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J. Hydrol. 509, 379–386.
Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. The University of Michigan Press, Ann Arbor, MI.
Jang, J.S., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst., Man, Cybern. 23 (3), 665–685.
Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., Xu, C.Y., 2015. Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters. J. Hydrol. 522, 326–338.
Jiang, C., Xiong, L., Guo, S., Xia, J., Xu, C.Y., 2017. A process-based insight into non-stationarity of the probability distribution of annual runoff. Water Resour. Res. 53 (5). https://doi.org/10.1002/2016WR019863.
Kasabov, N.K., Song, Q., 2002. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. IEEE Trans. Fuzzy Syst. 10 (2), 144–154.
Kasiviswanathan, K.S., He, J., Sudheer, K.P., Tay, J.H., 2016. Potential application of wavelet neural network ensemble to forecast streamflow for flood management. J. Hydrol. 536, 161–173.
Keskin, M.E., Taylan, D., Terzi, O., 2006. Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. Hydrol. Sci. J. 51 (4), 588–598.
Kumar, S., Tiwari, M.K., Chatterjee, C., Mishra, A., 2015. Reservoir inflow forecasting using ensemble models based on neural networks, wavelet analysis and bootstrap method. Water Resour. Manage. 29 (13), 4863–4883.
Lebrenz, H., Bárdossy, A., 2017. Estimation of the variogram using Kendall's Tau for a robust geostatistical interpolation. J. Hydrol. Eng. 22 (9), 04017038.
Li, L., Liu, P., Rheinheimer, D.E., Deng, C., Zhou, Y., 2014. Identifying explicit formulation of operating rules for multi-reservoir systems using genetic programming. Water Resour. Manage. 28 (6), 1545–1565.
Liu, Y.T., Lin, Y.Y., Wu, S.L., Chuang, C.H., Lin, C.T., 2016. Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. IEEE Trans. Neural Networks Learn. Syst. 27 (2), 347–360.
Lohani, A.K., Goel, N.K., Bhatia, K.K.S., 2014. Improving real time flood forecasting using fuzzy inference system. J. Hydrol. 509, 25–41.
Maidment, D.R., 1993. Handbook of Hydrology. McGraw-Hill, New York.
Mastorocostas, P.A., Theocharis, J.B., 2002. A recurrent fuzzy-neural model for dynamic system identification. IEEE Trans. Syst., Man, Cybern., Part B (Cybernetics) 32 (2), 176–190.
Méheust, D., Gangneux, J.P., Reponen, T., Wymer, L., Vesper, S., Le Cann, P., 2012. Correlation between Environmental Relative Moldiness Index (ERMI) values in French dwellings and other measures of fungal contamination. Sci. Total Environ. 438, 319–324.
Naghibi, S.A., Ahmadi, K., Daneshi, A., 2017. Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. Water Resour. Manage. 31 (9), 2761–2775.
Nanda, T., Sahoo, B., Beria, H., Chatterjee, C., 2016. A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products. J. Hydrol. 539, 57–73.
Nash, J.E., 1970. River flow forecasting through conceptual models, I: a discussion of principles. J. Hydrol. 10, 398–409.
Nguyen, S.D., Choi, S.B., Seo, T.I., 2018. Recurrent mechanism and impulse noise filter for establishing ANFIS. IEEE Trans. Fuzzy Syst. 26 (2), 985–997.
Nourani, V., Baghanam, A.H., Adamowski, J., Kisi, O., 2014. Applications of hybrid wavelet-Artificial Intelligence models in hydrology: a review. J. Hydrol. 514, 358–377.
Paudel, S., Elmitri, M., Couturier, S., Nguyen, P.H., Kamphuis, R., Lacarrière, B., Le Corre, O., 2017. A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. Energy Build. 138, 240–256.
Shen, M., Chen, J., Zhuan, M., Chen, H., Xu, C.Y., Xiong, L., 2018. Estimating uncertainty and its temporal variation related to global climate models in quantifying climate change impacts on hydrology. J. Hydrol. 556, 10–24.
Shoaib, M., Shamseldin, A.Y., Melville, B.W., Khan, M.M., 2016. A comparison between wavelet based static and dynamic neural network approaches for runoff prediction. J. Hydrol. 535, 211–225.
Taieb, S.B., Atiya, A.F., 2016. A bias and variance analysis for multistep-ahead time series forecasting. IEEE Trans. Neural Networks Learn. Syst. 27 (1), 62–76.
Talebizadeh, M., Moridnejad, A., 2011. Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models. Expert Syst. Appl. 38 (4), 4126–4135.
Tamura, H., Tanno, K., Tanaka, H., et al., 2008. Recurrent type ANFIS using local search technique for time series prediction. IEEE Asia Pacific Conference on Circuits and Systems, pp. 380–383.
Tan, Q.F., Lei, X.H., Wang, X., Wang, H., Wen, X., Ji, Y., Kang, A.Q., 2018. An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. J. Hydrol.
Taormina, R., Chau, K.W., Sivakumar, B., 2015. Neural network river forecasting through baseflow separation and binary-coded swarm optimization. J. Hydrol. 529, 1788–1797.
Tran, H.D., Muttil, N., Perera, B.J.C., 2016. Enhancing accuracy of autoregressive time series forecasting with input selection and wavelet transformation. J. Hydroinf. 18 (5), 791–802.
Tsai, M.J., Abrahart, R.J., Mount, N.J., Chang, F.J., 2014. Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan. Hydrol. Process. 28 (3), 1055–1070.
Wang, W.C., Chau, K.W., Cheng, C.T., Qiu, L., 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. J. Hydrol. 374 (3–4), 294–306.
Wang, W.C., Xu, D.M., Chau, K.W., Chen, S., 2013. Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. J. Hydroinf. 15 (4), 1377–1390.
Xiong, J.J., Zhang, G., 2018. Improved stability criterion for recurrent neural networks with time-varying delays. IEEE Trans. Neural Networks Learn. Syst. 26 (2), 1–5.
Yaseen, Z.M., El-Shafie, A., Jaafar, O., Afan, H.A., Sayl, K.N., 2015. Artificial intelligence based models for stream-flow forecasting: 2000–2015. J. Hydrol. 530, 829–844.
Yaseen, Z.M., Ebtehaj, I., Bonakdari, H., Deo, R.C., Mehr, A.D., Mohtar, W.H.M.W., Singh, V.P., 2017. Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model. J. Hydrol. 554, 263–276.
Yin, Z., Wen, X., Feng, Q., He, Z., Zou, S., Yang, L., 2017. Integrating genetic algorithm and support vector machine for modeling daily reference evapotranspiration in a semi-arid mountain area. Hydrol. Res. 48 (5), 1177–1191.
Zhang, D., Martinez, N., Lindholm, G., Ratnaweera, H., 2018a. Manage sewer in-line storage control using hydraulic model and recurrent neural network. Water Resour. Manage. 32 (6), 2079–2098.
Zhang, J., Morris, A.J., 1999. Recurrent neuro-fuzzy networks for nonlinear process modeling. IEEE Trans. Neural Networks 10 (2), 313–326.
Zhang, J., Zhu, Y., Zhang, X., Ye, M., Yang, J., 2018b. Developing a long short-term memory (LSTM) based model for predicting water table depth in agricultural areas. J. Hydrol. 561, 918–929.
Zhou, Y., Guo, S., 2013. Incorporating ecological requirement into multipurpose reservoir operating rule curves for adaptation to climate change. J. Hydrol. 498, 153–164.
Zhou, Y., Guo, S., Hong, X., Chang, F.J., 2017. Systematic impact assessment on inter-basin water transfer projects of the Hanjiang River Basin in China. J. Hydrol. 553, 584–595.
Zounemat-Kermani, M., Teshnehlab, M., 2008. Using adaptive neuro-fuzzy inference system for hydrological time series prediction. Appl. Soft Comput. 8 (2), 928–936.